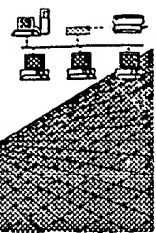


7200

BIOTECHNOLOGY  
SYSTEMS  
BRANCH



13

PK

## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number:

09/554,533

Source:

pat/09

Date Processed by STIC:

7/18/2003

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to:  
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7<sup>th</sup> Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202  
Or  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 04/24/2003

BEST AVAILABLE COPY

Raw Sequence Listing Error Summary

**ERROR DETECTED**

**SUGGESTED CORRECTION**

SERIAL NUMBER: 09/554,533

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1    Wrapped Nucleics  
    Wrapped Aminos      The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2    Invalid Line Length      The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3    Misaligned Amino  
    Numbering      The numbering under each 5<sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4    Non-ASCII      The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5    Variable Length      Sequence(s)   4   contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6    PatentIn 2.0  
    "bug"      A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s)           . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7    Skipped Sequences  
    (OLD RULES)      Sequence(s)            missing. If intentional, please insert the following lines for each skipped sequence:  
    (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
    (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)  
    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)  
    This sequence is intentionally skipped  
  
    Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8    Skipped Sequences  
    (NEW RULES)      Sequence(s)            missing. If intentional, please insert the following lines for each skipped sequence.  
    <210> sequence id number  
    <400> sequence id number  
    000
- 9    Use of n's or Xaa's  
    (NEW RULES)      Use of n's and/or Xaa's have been detected in the Sequence Listing.  
    Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.  
    In <220> to <223> section, please explain location of n or Xaa; and which residue n or Xaa represents.
- 10    Invalid <213>  
    Response      Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11    Use of <220>      Sequence(s)            missing the <220> "Feature" and associated numeric identifiers and responses.  
    Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.  
    (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12    PatentIn 2.0  
    "bug"      Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13    Misuse of n      n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH - Biotechnology Systems Branch - 08/21/2001

BEST AVAILABLE COPY



PCT09

## RAW SEQUENCE LISTING

DATE: 07/18/2003

PATENT APPLICATION: US/09/554,533

TIME: 13:08:11

Input Set : A:\EP.txt

Output Set: N:\CRF4\07182003\I554533.raw

4 <110> APPLICANT: BEELEY, NIGEL ROBERT ARNOLD  
 5 PRICKETT, KATHRYN S.  
 7 <120> TITLE OF INVENTION: NOVEL EXENDIN AGONIST COMPOUNDS  
 W--> 8 <130> FILE REFERENCE: 238/086 US (030639.0043.UTL2)  
 10 <140> CURRENT APPLICATION NUMBER: 09/554,533  
 11 <141> CURRENT FILING DATE: 1998-11-13  
 13 <150> PRIOR APPLICATION NUMBER: PCT/US98/24210  
 14 <151> PRIOR FILING DATE: 1998-11-13  
 16 <150> PRIOR APPLICATION NUMBER: US 60/065,442  
 17 <151> PRIOR FILING DATE: 1997-11-14  
 19 <160> NUMBER OF SEQ ID NOS: 74  
 21 <170> SOFTWARE: FastSEQ for Windows Version 3.0

## ERRORED SEQUENCES

68 <210> SEQ ID NO: 3  
 69 <211> LENGTH: 30  
 70 <212> TYPE: PRT  
 71 <213> ORGANISM: Homo sapien  
 W--> 72 <220> FEATURE:  
 73 <221> NAME/KEY: AMIDATION  
 74 <222> LOCATION: (30)...(30)  
 75 <223> OTHER INFORMATION: amidated Arg (Argininamide)  
 77 <400> SEQUENCE: 3  
 79 His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly  
 80 1 5 10 15  
 82 Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg  
 E--> 83 ~~25~~ 20 30 25 30  
 85 <210> SEQ ID NO: 4  
 86 <211> LENGTH: 29  
 87 <212> TYPE: PRT  
 88 <213> ORGANISM: Artificial Sequence  
 90 <220> FEATURE:  
 91 <223> OTHER INFORMATION: artificially synthesized sequence of novel exendin agonist  
 92 compound  
 94 <220> FEATURE:  
 95 <221> NAME/KEY: VARIANT  
 96 <222> LOCATION: (1)...(7)  
 97 <223> OTHER INFORMATION: Xaa in position 1 is His, Arg or Tyr; Xaa in position 2 is  
 98 Ser, Gly, Ala or Thr; Xaa in position 3 is Asp or Glu;  
 99 Xaa in position 5 is Ala or Thr; Xaa in position 6 is Ala,  
 100 Phe, Tyr or naphthylalanine; Xaa in position 7 is Thr or Ser;

pp 1-3,5  
 Does Not Comply  
 Corrected Diskette Needed

misaligned  
 amino acid  
 20 number -  
 see item 3  
 on Err Summary  
 sheet

## RAW SEQUENCE LISTING

DATE: 07/18/2003

PATENT APPLICATION: US/09/554,533

TIME: 13:08:11

Input Set : A:\EP.txt

Output Set: N:\CRF4\07182003\I554533.raw

102 <220> FEATURE:  
 103 <221> NAME/KEY: VARIANT  
 104 <222> LOCATION: (8)...(13)  
 105 <223> OTHER INFORMATION: Xaa in position 8 is Ala, Ser or Thr; Xaa in position 9 is  
 106 Asp or Glu; Xaa in position 10 is Ala, Leu, Ile, Val, pentyl-  
 107 glycine or Met; Xaa in position 11 is Ala or Ser; Xaa in  
 108 position 12 is Ala or Lys; Xaa in position 13 is Ala or Gln;  
 110 <220> FEATURE:  
 111 <221> NAME/KEY: VARIANT  
 112 <222> LOCATION: (14)...(20)  
 113 <223> OTHER INFORMATION: Xaa in position 14 is Ala, Leu, Ile, pentylglycine, Val or  
 114 Met; Xaa in position 15 is Ala or Glu; Xaa in position 16 is  
 115 Ala or Glu; Xaa in position 17 is Ala or Glu; Xaa in position  
 116 19 is Ala or Val; Xaa in position 20 is Ala or Arg;  
 118 <220> FEATURE:  
 119 <221> NAME/KEY: VARIANT  
 120 <222> LOCATION: (21)...(24)  
 121 <223> OTHER INFORMATION: Xaa in position 21 is Ala or Leu; Xaa in position 22 is Phe,  
 122 Tyr or naphthylalanine; Xaa in position 23 is Ile, Val, Leu,  
 123 pentylglycine, tert-butylglycine or Met; Xaa in position 24  
 124 is Ala, Glu or Asp;  
 126 <220> FEATURE:  
 127 <221> NAME/KEY: VARIANT  
 128 <222> LOCATION: (25)...(28)  
 129 <223> OTHER INFORMATION: Xaa in position 25 is Ala, Trp, Phe, Tyr or naphthylalanine;  
 130 Xaa in position 26 is Ala or Leu; Xaa in position 27 is Ala  
 131 or Lys; Xaa in position 28 is Ala or Asn;  
 133 <220> FEATURE:  
 134 <221> NAME/KEY: VARIANT  
 135 <222> LOCATION: (29)...(29)  
 136 <223> OTHER INFORMATION: Xaa in position 29 is -OH, -NH<sub>2</sub>, Gly-Z2; Gly Gly-Z2; Gly Gly  
 137 Xaa31-Z2; Gly Gly Xaa31 Ser-Z2; Gly Gly Xaa31 Ser Ser-Z2; Gly Gly  
 138 Xaa31 Ser Ser Gly-Z2; Gly Gly Xaa31 Ser Ser Gly Ala-Z2; Gly Gly  
 139 Xaa31 Ser Ser Gly Ala Xaa36-Z2;  
 141 <220> FEATURE:  
 142 <221> NAME/KEY: VARIANT  
 143 <222> LOCATION: (29)...(29)  
 144 <223> OTHER INFORMATION: Gly Gly Xaa31 Ser Ser Gly Ala Xaa36 Xaa37-Z2; or Gly Gly  
 145 Ser Ser Gly Ala Xaa36 Xaa37 Xaa38-Z2;  
 147 <220> FEATURE:  
 148 <221> NAME/KEY: VARIANT  
 149 <222> LOCATION: (29)...(29)  
 150 <223> OTHER INFORMATION: where Xaa31, Xaa36, Xaa37 and Xaa38 are independently Pro,  
 151 homoproline, 3Hyp, 4Hyp, thioproline, N-alkylglycine,  
 152 N-alkylpentylglycine or N-alkylalanine; and Z2 is -OH or -NH<sub>2</sub>;  
 154 <220> FEATURE:  
 155 <221> NAME/KEY: VARIANT  
 156 <222> LOCATION: (3)...(28)  
 157 <223> OTHER INFORMATION: provided that no more than three of Xaa in positions 3, 5,

*Xaa can only represent a single amino acid, nothing else*

*see item 5 on Erra summary sheet*

*Only 29 amino acids in sequence 4*

## RAW SEQUENCE LISTING

DATE: 07/18/2003

PATENT APPLICATION: US/09/554,533

TIME: 13:08:11

Input Set : A:\EP.txt

Output Set: N:\CRF4\07182003\I554533.raw

158 8, 10, 11, 12, 13, 14, 15, 16, 17, 19, 20, 21, 24, 25, 26,  
 159 27 and 28 are Ala.  
 161 <400> SEQUENCE: 4  
 162 Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa 1  
 E--> 163 5 10 15  
 W--> 164 Xaa Ala Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
 E--> 165 20 25  
 168 <210> SEQ ID NO: 5  
 169 <211> LENGTH: 30  
 170 <212> TYPE: PRT  
 171 <213> ORGANISM: Artificial Sequence  
 173 <220> FEATURE:  
 174 <223> OTHER INFORMATION: artificially synthesized sequence of novel exendin agonist  
 175 compound  
 177 <220> FEATURE:  
 178 <221> NAME/KEY: AMIDATION  
 179 <222> LOCATION: (30)...(30)  
 180 <223> OTHER INFORMATION: amidated Gly (Glycinamide)  
 182 <400> SEQUENCE: 5  
 183 His Gly Glu Gly Thr Phe Thr Ser Asp Leu Ser Lys Gln Met Glu Glu 1  
 E--> 184 5 10 15  
 186 Glu Ala Val Arg Leu Phe Ile Glu Trp Leu Lys Asn Gly Gly  
 E--> 187 25 30

*insert a  
hard return  
and re-number  
the amino  
acids  
under every  
5 amino acids*

*insert a  
hard return  
and  
re-number  
amino acids*

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/554,533

DATE: 07/18/2003  
TIME: 13:08:13

Input Set : A:\EP.txt  
Output Set: N:\CRF4\07182003\I554533.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:3; Line(s) 82  
Seq#:4; Line(s) 162  
Seq#:5; Line(s) 183,186

## VARIABLE LOCATION SUMMARY

DATE: 07/18/2003

PATENT APPLICATION: US/09/554,533

TIME: 13:08:13

Input Set : A:\EP.txt

Output Set: N:\CRF4\07182003\I554533.raw

Use of n's or Xaa's (NEW RULES) :

Use of n's and/or Xaa's have been detected in the Sequence Listing.

Use of &lt;220&gt; to &lt;223&gt; is MANDATORY if n's or Xaa's are present.

in &lt;220&gt; to &lt;223&gt; section, please explain location of n or Xaa, and which residue n or Xaa represents.

Seq#:4; Xaa Pos. 17,19,20,21,22,23,24,25,26,27,28,29

Seq#:47; Xaa Pos. 31,36,37,38

Seq#:48; Xaa Pos. 36,37,38

Seq#:51; Xaa Pos. 31,36,37

Seq#:52; Xaa Pos. 31,36

Seq#:55; Xaa Pos. 6

Seq#:59; Xaa Pos. 10

Seq#:60; Xaa Pos. 22

Seq#:61; Xaa Pos. 23

Seq#:65; Xaa Pos. 31,36,37

Seq#:66; Xaa Pos. 1,2,3,5,6,7,8,9,10,11,12,13,14,15,16,17,19,20,21,22,23,24

Seq#:66; Xaa Pos. 25,26,27,28

## VERIFICATION SUMMARY

DATE: 07/18/2003

PATENT APPLICATION: US/09/554,533

TIME: 13:08:13

Input Set : A:\EP.txt

Output Set: N:\CRF4\07182003\I554533.raw

L:8 M:283 W: Missing Blank Line separator, <130> field identifier  
L:72 M:283 W: Missing Blank Line separator, <220> field identifier  
L:83 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:3  
L:83 M:252 E: No. of Seq. differs, <211> LENGTH:Input:30 Found:16 SEQ:3  
L:163 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:4  
L:164 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0  
M:332 Repeated in SeqNo=4  
L:165 M:252 E: No. of Seq. differs, <211> LENGTH:Input:29 Found:13 SEQ:4  
L:184 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:5  
M:332 Repeated in SeqNo=5  
L:187 M:301 E: (44) No Sequence Data was Shown, SEQ ID:5  
L:187 M:252 E: No. of Seq. differs, <211> LENGTH:Input:30 Found:0 SEQ:5  
L:1186 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:47 after pos.:16  
M:341 Repeated in SeqNo=47  
L:1218 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:48 after pos.:32  
L:1302 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:51 after pos.:16  
M:341 Repeated in SeqNo=51  
L:1331 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52 after pos.:16  
M:341 Repeated in SeqNo=52  
L:1405 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:55 after pos.:0  
L:1500 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:59 after pos.:0  
L:1529 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:60 after pos.:16  
L:1546 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:61  
L:1553 M:258 W: Mandatory Feature missing, <222> Tag not found for SEQ ID#:61  
L:1553 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:61 after pos.:16  
L:1615 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:64  
L:1640 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:65  
L:1647 M:258 W: Mandatory Feature missing, <222> Tag not found for SEQ ID#:65  
L:1647 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:65 after pos.:16  
M:341 Repeated in SeqNo=65  
L:1743 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:66 after pos.:0  
M:341 Repeated in SeqNo=66